Machine Learning and Artificial Intelligence Applications in Metallurgical Engineering: A Comprehensive Review

James Joy P
Andhra University, Visakhapatnam 530003, Andhra Pradesh, India
doi.org/10.64643/IJIRTV13I1-203942-459

***Abstract*—Over the past decade, machine learning (ML) and artificial intelligence (AI) have moved from peripheral curiosities to working tools across metallurgical engineering, reshaping how researchers and plant engineers approach materials discovery, process optimisation, and quality assurance. This review surveys that shift. We examine ML methods (supervised and unsupervised learning, deep neural networks, and reinforcement learning) as they have been applied to phase diagram prediction, mechanical property forecasting, steelmaking process control, corrosion modelling, heat treatment, additive manufacturing, and non-destructive evaluation. Throughout, our interest is in how traditionally empirical, experience-driven practice is giving way to data-driven, predictive frameworks. Drawing on 128 publications from 2015 to 2025, we map the prevailing algorithmic trends, point to benchmark datasets, and weigh the practical obstacles to deployment: limited data, weak physical interpretability, and uneasy integration with computational thermodynamics (CALPHAD). We also identify open research gaps and sketch a roadmap that includes physics-informed neural networks (PINNs), graph neural networks for microstructure representation, and federated learning for industrial data sharing. The aim is to provide a single point of reference for metallurgical researchers, process engineers, and educators moving into data-driven materials science.**

***Index Terms*— Machine Learning; Artificial Intelligence; Metallurgical Engineering; Materials Informatics; Neural Networks; Process Optimisation; Mechanical Properties; Phase Prediction**

I. INTRODUCTION

The global metals and materials industry is in the middle of a digital shift, pulled along by two pressures that are easy to state and hard to satisfy: shorten product development cycles, and reduce the environmental footprint of processes that remain stubbornly energy-intensive. The conventional toolkit for alloy design, process optimisation, and quality control consists of empirical correlations, thermodynamic equilibrium calculations, and the heuristics passed down through decades of plant practice. It has carried the field a long way. It has produced advanced high-strength steels (AHSS), nickel superalloys, titanium biomedical implants, and much else. But the combinatorial complexity of modern multicomponent alloys has now outgrown what purely experimental campaigns or classical computational methods can reasonably explore [1, 2].

Machine learning and AI, broadly understood as computational frameworks that learn patterns from data rather than following hand-coded rules, offer a different route. Three developments have made the shift both timely and practical: the rapid growth of materials databases (ICSD, Materials Project, AFLOW, NIST, Citrine); falling computational cost, which has opened the door to high-throughput density functional theory (DFT) screening; and the maturity of open-source frameworks such as TensorFlow, PyTorch, and scikit-learn, which have placed the algorithms within reach of any reasonably equipped lab [3, 4]. Metallurgy carries a fourth advantage that is often overlooked: decades of operating data sitting in integrated steel mills, foundries, and heat treatment shops. That archive has been underused, and ML methods are unusually well suited to making something of it.

The Materials Genome Initiative (MGI), launched in 2011, gave materials informatics formal standing as a discipline and set a deliberately ambitious target: halve the time and cost of materials discovery and deployment [5]. Influential papers followed. Raccuglia et al. [6] used ML to guide synthesis planning, and Ward et al. [7] introduced the matminer feature library for property prediction. Together they established a working proof of concept for ML in

inorganic materials. Metallurgy was slower to take up these methods, partly because microstructure–process–property relationships are unforgivingly tangled, and partly because much of the relevant data is locked inside proprietary industrial archives. Even so, the publication rate has climbed sharply since 2018 [8].

For all this growth, the literature still lacks a review that spans the metallurgical sub-disciplines, from blast furnace ironmaking to additive manufacturing, while also taking deployment readiness seriously. Existing reviews tend to stay within their domains: Liu et al. [9] on ML for steel properties, Schmidt et al. [10] on materials ML in general, and Dimiduk et al. [11] on AI for microstructure analysis. The present paper attempts a wider cross-domain view, with one further motivation: the Indian metallurgical sector (RINL, SAIL, JSW, TATA Steel) is at a point where Industry 4.0 practices are moving from pilot projects to operational reality, and a consolidated account of what is and is not working has practical value here.

The review has four objectives. First, to map where ML and AI are actually being applied across the major domains of metallurgical engineering. Second, to ask whether the algorithmic choices being made are well matched to the kinds of datasets metallurgy produces. Third, to surface the recurring difficulties (data quality, interpretability, physical consistency) that limit what these models can do today. And fourth, to set out a roadmap for combining ML with physics-based models in a way that is both accurate and scientifically defensible.

II. LITERATURE SEARCH METHODOLOGY

We searched Scopus, Web of Science, Google Scholar, and the Materials Project database for publications between January 2015 and March 2025. The Boolean queries combined terms such as "machine learning + metallurgy", "neural network + steel properties", "artificial intelligence + alloy design", and "deep learning + microstructure", along with several domain-specific variants. The initial pool came to 324 articles. After removing duplicates and applying our inclusion criteria (peer-reviewed, English language, and a direct ML or AI application to a metallurgical problem), 128 articles remained for detailed review. Table 1 shows how they distribute across sub-domains.

Table 1. Distribution of reviewed publications across metallurgical sub-domains (2015–2025).

Sub-Domain	No. of Papers	Primary ML Methods	Key Applications
Mechanical Property Prediction	28	ANN, RF, SVR, GBM	UTS, yield strength, hardness, fatigue life
Phase Diagram / CALPHAD-ML	18	GNNs, CALPHAD-ML hybrid	Phase boundary prediction, solidification
Steelmaking Process Control	22	LSTM, DNN, RL	BOF endpoint, BF control, ladle treatment
Corrosion Behaviour	14	RF, SVM, ANN	Corrosion rate, pitting susceptibility
Heat Treatment Optimisation	12	BO, DNN, GA-ANN	Temperature, time, quench medium selection
Additive Manufacturing	16	CNN, DNN, Bayesian Opt.	Process parameter optimisation, porosity
Microstructure Analysis	10	CNN, U-Net, Transfer Learning	Grain segmentation, phase identification
NDT / Defect Detection	8	CNN, SVM, Autoencoder	Crack detection, inclusion identification
TOTAL	128		

III. MACHINE LEARNING FUNDAMENTALS RELEVANT TO METALLURGY

A. Overview of Key Algorithms

Supervised learning, which trains on labelled input–output pairs, accounts for most of the ML work in metallurgy, simply because property-measurement databases provide the labels. Artificial neural networks (ANNs), together with their deeper variants, have been the workhorse since Bhadeshia's 1999 paper introduced the metallurgical community to neural network models for steel properties [12]. Random forests (RF), the decision-tree ensemble introduced by Breiman [13], have earned their popularity for two reasons: they resist overfitting, and the feature importance scores they produce are themselves informative. Knowing which compositional or processing variable matters most is often half the scientific question. Gradient boosting machines (GBMs), notably XGBoost and LightGBM, have consistently led the accuracy rankings on tabular metallurgical datasets in benchmark comparisons [14]. Support vector regression (SVR) with radial basis function kernels holds up surprisingly well in the

small-data regime (n < 200) that experimental metallurgical studies often inhabit. Gaussian process regression (GPR), a Bayesian non-parametric method, is a particular favourite for materials discovery because every prediction comes with an uncertainty estimate. That is exactly what is needed to drive active learning and Bayesian optimisation (BO) loops in guided experimentation [15]. For image-based work, including microstructure analysis, grain boundary segmentation, and defect detection in SEM and EBSD datasets, convolutional neural networks (CNNs) have become the default [16].

B. Feature Engineering in Metallurgical Datasets

No ML model in metallurgy will outperform its input features. Three families of descriptors recur. The first is compositional: elemental fractions, weighted-average atomic properties such as atomic radius, electronegativity, valence electron count, and melting point, together with thermodynamically derived terms (mixing enthalpy, ΔH_{mix}, and mixing entropy, ΔS_{mix}) that are essential for any high-entropy alloy (HEA) study [17]. The second comes from structure: lattice parameters and peak intensities from X-ray diffraction, alongside microstructural quantities such as grain size, phase fraction, and precipitate morphology. The third captures processing history (cold reduction ratio, authentication temperature, quench rate, tempering temperature), and is the descriptor class most directly under the operator's control [18]. For practitioners, the matminer Python library provides a ready-made featurisation pipeline for compositional, structural, and site-based descriptors drawn from the Materials Project database [7].

IV. ML FOR MECHANICAL PROPERTY PREDICTION

Predicting mechanical properties (tensile strength, yield strength, elongation, hardness, fatigue life, fracture toughness) from composition and processing parameters is the most heavily studied ML application in metallurgy [19, 20]. The classical alternatives are familiar: phenomenological models such as Hall–Petch for grain size dependence and the Hollomon equation for strain hardening, or CALPHAD-informed physical models. Both demand extensive characterisation, and neither generalises easily across alloy families.

Liu et al. [9] assembled a dataset of 1,319 low-alloy steels with composition and heat treatment parameters and trained RF, SVR, and ANN models to predict tensile strength and elongation. RF reached RMSE = 31 MPa ($R^2 = 0.96$) for tensile strength, comfortably outperforming the conventional carbon equivalent formula. Just as important, the RF feature importance picked out carbon content, manganese, and authentication temperature as the dominant variables. This is what any metallurgist would expect, and it provides exactly the kind of sanity check the ML approach needs [9].

Ma et al. [21] built a deep neural network (DNN) for nickel-based superalloy properties using a dataset of 827 compositions. Feeding the model both alloy chemistry and aging treatment parameters, they reached mean absolute percentage errors (MAPE) below 5% for creep rupture strength, a property that is notoriously hard to model analytically because it depends on microstructural features that composition alone simply cannot describe. In a separate study, transfer learning from a large steel dataset to a much smaller titanium alloy set (n = 85) lifted prediction accuracy by 23% over training on the titanium data alone [22], which speaks directly to one of the field's perennial problems: tiny datasets.

High entropy alloys (HEAs), with five or more principal elements in play, have made ML something close to indispensable. Huang et al. [23] screened more than two million candidate HEA compositions, pairing thermodynamic stability criteria with RF classification for hardness, and surfaced 14 novel candidates with HV > 600. Zhang et al. [24] used Bayesian optimisation to explore the CoCrFeMnNi–Al–Ti space and arrived at alloys 40% stronger by specific strength than the Cantor alloy in just 12 experimental iterations. It is a striking illustration of how much the experimental burden can be reduced when ML guides the search.

Fatigue life prediction is fundamentally a statistical problem, since scatter is large and surface and microstructural variability matter enormously, so it is natural ground for probabilistic ML. Agrawal and Choudhary [25] used an ensemble of 1,000 bootstrapped neural networks to construct probabilistic S–N curves for steel components, returning confidence intervals alongside mean fatigue life. Compared with the deterministic single-number

output of conventional models, that is a meaningful step forward.

V. ML IN STEELMAKING AND IRONMAKING PROCESS CONTROL

Integrated steelmaking is a chain of tightly coupled unit operations (blast furnace, basic oxygen furnace or BOF, ladle metallurgy, continuous casting), and each one generates the kind of data ML thrives on: time-series sensor readings, chemical analyses, temperature measurements, all of it at high frequency. The potential value is large. In a 5 MTPA plant, a one percentage point improvement in BOF hit rate (hitting target carbon and temperature simultaneously) translates into tens of millions of dollars in annual savings through reduced reblowing, fewer sampling delays, and less energy wasted [26].

A. Blast Furnace Control

The blast furnace is a non-linear, distributed-parameter system whose internal state cannot be measured directly, and that description captures both the challenge and the appeal of treating it with ML. Models have been used to predict hot metal silicon content, which acts as a proxy for furnace thermal state, and to forecast hanging and slipping events and raceway conditions. Yuan et al. [27] deployed an LSTM (long short-term memory) recurrent network on a rolling 24-hour window of 48 process variables to predict silicon content two hours ahead, hitting ±0.05% Si on 87% of heats. That kind of lead time lets operators adjust burden and blast before the furnace drifts into abnormal operation, rather than reacting after it has.

B. BOF Endpoint Prediction

BOF endpoint control, which means predicting final carbon content and temperature from sub-lance measurements and gas emissions, is one of the longest-running ML applications in steelmaking. Tang et al. [28] compared ANN, SVM, RF, and XGBoost on 3,842 BOF heats from a Chinese integrated steelwork. XGBoost came out ahead on endpoint carbon, with a hit rate of 91.3% within ±0.015% C, and the dominant predictors were initial hot metal chemistry, scrap ratio, oxygen consumption, and the evolution of the off-gas CO/CO_2 ratio. The more interesting finding was operational: feeding the off-gas analysis into the model in real time improved the hit rate by 8 percentage points over models using only static inputs.

C. Continuous Casting Quality

Breakout prediction, the task of catching imminent shell ruptures in the continuous casting mould before they become costly production stoppages, has been an ML target since the 1990s using thermocouple-array data. Recent deep learning work has cut false alarm rates sharply. Zhang et al. [29] applied a 2D CNN to the spatiotemporal thermocouple pattern from a 220 × 220 mm billet mould and predicted breakouts 90 seconds in advance, with a 97% true positive rate and fewer than 0.2 false alarms per week. Both numbers represent a clear improvement over the classical thermal mapping approach.

VI. ML-ENHANCED PHASE PREDICTION AND CALPHAD INTEGRATION CALPHAD

(Calculation of Phase Diagrams) remains the gold standard for phase equilibria in multicomponent systems, and rightly so: where the thermodynamic database is well parameterised, it is hard to beat. The difficulty is at the boundaries. CALPHAD depends on extensive experimental validation, and it falters in compositional regions where that validation does not yet exist, which happens to be exactly where ML can offer the most help [30].

Graph neural networks (GNNs) are a natural fit for phase prediction because they encode crystal structures directly as atomic graphs. The pioneering example is the Crystal Graph Convolutional Neural Network (CGCNN) of Xie and Grossman [31], which predicted formation energies, bandgaps, and bulk moduli for more than 29,000 materials at DFT-comparable accuracy in milliseconds. This opened the door to high-throughput screening workflows that would be infeasible with direct DFT. For HEAs, Zhou et al. [32] trained a classifier on 401 experimentally verified compositions to predict phase formation (single FCC, single BCC, mixed, or amorphous) and reached 85% accuracy using only thermodynamic parameters (ΔH_{mix}, ΔS_{mix}, δ, and VEC) as features.

Hybrid CALPHAD–ML frameworks take this further. By using ML as a surrogate for the CALPHAD calculation itself, typically at around three orders of magnitude lower computational cost, they have been

demonstrated for solidification pathway prediction [33] and TTT diagram construction [34]. The use case where this matters most is finite element solidification modelling, where each step requires a thermodynamic call and those calls have historically been the bottleneck.

VII. ML IN CORROSION BEHAVIOUR MODELLING

Corrosion costs the world economy something on the order of 3.4% of global GDP each year [35], so improvements in corrosion prediction carry real economic weight. The classical toolset (Tafel kinetics, the Stern–Geary equation, equivalent-circuit fits to electrochemical impedance spectra) is excellent at explaining mechanisms, but it has never been particularly good at predicting long-term service behaviour in complex environments. ML offers a complementary, data-driven angle.

Sun et al. [36] collected 1,127 electrochemical measurements on stainless steels in chloride environments and trained RF and gradient boosting models to predict pitting potential (E_pit) from alloy composition and electrolyte chemistry (Cl^- concentration, pH, temperature). The RF model reached $R^2 = 0.91$ on the test set, identifying chromium equivalent (Cr + 3Mo + 0.5Si) and chloride activity as the most influential features. The encouraging part was that the model reproduced a well-known experimental trend on its own: molybdenum additions above 2 wt% shifted E_pit by +150 to +200 mV in 3.5% NaCl. That kind of agreement is what makes the model physically interpretable, not merely accurate.

For atmospheric corrosion of infrastructure steels, deep learning models that combine meteorological data (temperature, humidity, pollution indices) with exposure time and steel composition have been trained on the ISO CORRAG programme, a 12-year multi-site atmospheric exposure study. Wang et al. [37] achieved RMSE < 8 μm/year on 8-year corrosion depth predictions, beating the ISO 9224 power-law model by 35% for inland industrial environments.

Hydrogen embrittlement (HE) susceptibility prediction is an emerging area with direct industrial stakes, especially for high-strength pipeline and automotive steels. Jain et al. [38] trained an ANN on 283 literature data points to predict the HE indexes from steel composition, hydrogen concentration, and microstructural parameters. What the model surfaced was the strongly non-linear interplay between carbon content and prior austenite grain size, an interaction that the HELP (Hydrogen Enhanced Localised Plasticity) mechanism predicts qualitatively, but that had not previously been quantified empirically in this form.

VIII. ML IN HEAT TREATMENT OPTIMISATION

Heat treatment design, which involves choosing austenitisation temperature and time, quench medium, and tempering temperature and time so as to land target mechanical properties, is traditionally a matter of reading TTT/CCT diagrams in light of accumulated empirical experience. ML has been used in both directions: as a forward model that predicts properties from a given heat treatment schedule, and as an inverse design tool that recommends a schedule for a desired property target.

Popova et al. [39] paired Bayesian optimisation with a DNN surrogate to design a heat treatment for a nickel superalloy, targeting tensile strength above 1200 MPa and elongation above 12% simultaneously. The BO algorithm converged on an optimal ageing sequence in 18 experimental iterations, against an estimated 200+ that a conventional design of experiments would have required, a 90% reduction in experimental effort. The optimum it identified, a two-step ageing at 1079 °C/1 h followed by 760 °C/16 h, had not been reported in the open literature.

For medium carbon steels in automotive crankshafts, Wu et al. [40] trained an ensemble of 500 bootstrapped RF models on 1,243 heat treatment records from a production facility to build a probabilistic property predictor. The system was deployed as a real-time recommendation engine, and over a six-month trial the proportion of out-of-specification components fell from 3.2% to 0.8%, a result that translates directly into cost savings and better product quality.

IX. ML IN ADDITIVE MANUFACTURING OF METALS

Metal additive manufacturing, particularly selective laser melting (SLM) and electron beam melting (EBM), is distinguished by how unforgivingly process-sensitive it is. The interplay of laser power, scan speed, hatch spacing, layer thickness, and

atmosphere governs porosity, residual stress, microstructure, and ultimately mechanical performance [41]. The four-dimensional parameter space is large enough that exhaustive experimental optimisation is essentially out of reach, which makes ML-guided parameter selection a natural fit.

Tapia et al. [42] used Gaussian process regression with active learning to optimise SLM parameters for 316L stainless steel, with minimum porosity as the objective. Starting from just 12 initial experiments, the GPR–BO framework arrived at parameter combinations producing less than 0.1% porosity after only 8 additional runs. By comparison, Taguchi or full-factorial designs typically demand 50 to 100 experiments. The GPR uncertainty estimates were doing the real work here, steering experiments toward the most informative regions of parameter space.

In-situ quality monitoring is an even more direct application. CNN-based classification of melt pool images captured by high-speed cameras can flag porosity and lack-of-fusion defects in real time during SLM builds. Ye et al. [43] trained a ResNet-50 CNN on 45,000 labelled melt pool images from Ti-6Al-4V SLM experiments, reaching 94.7% accuracy on binary defective/non-defective classification with an inference time of 3 ms per image, fast enough for real-time feedback control at standard scan speeds.

X. AI-DRIVEN MICROSTRUCTURE CHARACTERISATION

Microstructural quantification (grain size measurement, phase fraction determination, precipitate morphology analysis) has traditionally meant slow, manual image analysis. CNN-based semantic segmentation has largely automated these tasks, with a consistency and speed that human operators cannot match [44].

DeCost and Holm [45] were the first to apply deep CNNs to steel microstructure classification, training a model on 148 micrographs from the Ultra High Carbon Steel (UHCS) dataset to distinguish 12 microstructural classes (pearlite, spheroidite, Widmanstätten, martensite, and others) with 90% top-1 accuracy. Azimi et al. [46] followed up using fully convolutional networks with an encoder–decoder architecture for grain boundary segmentation in austenitic steel EBSD maps, achieving grain size measurement errors below 3% against the manual ASTM E112 standard and reducing analysis time from hours to seconds.

In nickel superalloys, TEM image analysis with instance segmentation (Mask R-CNN) has been used to quantify γ′ precipitate size, shape, and volume fraction automatically across hundreds of images. The point is not just speed: that volume of analysis allows statistically robust microstructure–property correlations that were not previously practical, simply because the manual workload was too great [47].

XI. CHALLENGES AND LIMITATIONS

A. Data Scarcity and Quality

Metallurgical ML works with datasets that are orders of magnitude smaller than those routinely used in computer vision or NLP, typically n = 100 to 10,000, against the millions in ImageNet. Experimental data collection is expensive and slow, and is often gated by equipment availability. Industrial data from steel plants compounds the problem: it is commercially sensitive, and academic researchers rarely get access. The standard mitigation strategies are transferring learning from compositionally similar alloy systems [22], physics-informed data augmentation, and active learning frameworks that try to extract the maximum information from each experiment [15].

B. Physical Interpretability

Black-box models, and DNNs in particular, that cannot be tied back to physical mechanisms attract justified scepticism from materials scientists, and they are difficult to defend in safety-critical applications. Explainability tools have helped here. SHAP (SHapley Additive exPlanations) values, LIME, and integrated gradients have all been applied to materials science models to extract feature importances and to surface physically meaningful trends [48]. The deeper response, and probably the more promising one in the long run, is to build physical consistency into the model itself. Physics-informed neural networks (PINNs), which embed governing equations such as the heat equation, Fick's law of diffusion, or the Navier–Stokes equations as soft constraints in the loss function, represent the most active line of work along these lines [49].

C. Extrapolation Beyond Training Domain

Models trained on existing alloy databases tend to do poorly when asked to extrapolate to genuinely novel

compositions or processing windows outside their training distribution, which is, awkwardly, exactly where the highest-value discoveries are likely to be made. Uncertainty quantification, whether through Bayesian methods, deep ensembles, or Monte Carlo Dropout, is essential as a way of flagging when a prediction should not be trusted [50]. Applicability domain analysis ought to be a standard part of any metallurgical ML workflow, and is currently not.

XII. FUTURE DIRECTIONS AND RESEARCH ROADMAP

Drawing on the review above, several research directions stand out as priorities for the metallurgical ML community:

(1) Physics-informed ML. Embedding thermodynamic constraints (Gibbs free energy minimisation, the lever rule, constitutional supercooling criteria) as hard or soft constraints during neural network training should improve extrapolation reliability and reduce data requirements at the same time. Coupling PINNs with CALPHAD databases is an immediately actionable starting point.

(2) Foundation models for materials. Large pre-trained models, analogous to GPT for language but trained on millions of DFT calculations and experimental records, could act as feature extractors that need only modest fine-tuning for any specific metallurgical task. The early work on MatBERT [51] suggests there is real potential here.

(3) Multi-fidelity learning. Cheap, abundant low-fidelity data (CALPHAD calculations, empirical correlations) can be combined with sparse, expensive high-fidelity data such as careful experiments or high-accuracy DFT, through multi-fidelity Gaussian processes or co-kriging. The payoff is better data efficiency at the same overall cost.

(4) Digital twins for metallurgical plants. ML models embedded in real-time process control, fed by IoT sensor networks attached to blast furnaces, BOFs, and continuous casting machines, are the obvious next step toward what is now being called "intelligent steelmaking". RINL Visakhapatnam and Tata Steel's Jamshedpur plant are both running pilot implementations along these lines.

(5) Federated learning for industrial data. The commercial sensitivity of plant data can be partly resolved through federated learning, where models are trained across multiple plant datasets without raw data ever leaving each site. This makes collaborative model development possible without compromising commercial confidentiality.

(6) Sustainable metallurgy. ML-driven optimisation of hydrogen-based direct reduction (H-DRI), electrolytic ironmaking, and carbon capture in steel plants is an urgent priority in light of global decarbonisation targets. The strongly non-linear thermodynamics and kinetics of these newer processes are particularly well suited to ML-assisted optimisation.

XIII. CONCLUSIONS

This review has surveyed 128 publications on the use of ML and AI across the breadth of metallurgical engineering, from property prediction and phase diagram calculation to process control, corrosion modelling, heat treatment, additive manufacturing, and microstructure analysis. A few conclusions follow.

(i) Across every domain we examined, ML algorithms (random forests, gradient boosting, deep neural networks, and CNNs in particular) have delivered statistically significant improvements in predictive accuracy over the classical metallurgical models they are competing against.

(ii) Hybrid approaches that combine ML with CALPHAD thermodynamics, and physics-informed neural networks that incorporate constitutive equations, offer the most promising route to models that are both physically interpretable and capable of extrapolating beyond their training data.

(iii) Data scarcity remains the central challenge for metallurgical ML. Transfer learning, active learning, and multi-fidelity approaches are the mitigation strategies with the strongest track record so far.

(iv) The deployment of ML in industrial steelmaking (blast furnace control, BOF endpoint prediction, breakout prediction in continuous casting) has already shown tangible techno-economic benefits at scale. These are no longer laboratory demonstrations.

(v) The high-impact directions for the next several years are physics-informed neural networks, foundation models for materials, digital twins for integrated steelmaking, and ML-optimised green steel processes. Each is active, none is mature, and all reward further work.

With global steel production targeting net-zero emissions by 2050, and demand for advanced materials intensifying across aerospace, automotive, energy, and biomedical sectors, the case for integrating ML and AI into metallurgical engineering practice no longer rests on novelty. It rests on need. For the field to make full use of what has been built so far, the metallurgical engineering community will have to develop competences that bridge materials science, process engineering, and data science, and to do so without abandoning the physical intuition that gives the discipline its character.

REFERENCES

[1] Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation," *APL Materials*, vol. 1, no. 1, Art. no. 011002, 2013.

[2] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The high-throughput highway to computational materials design," *Nature Materials*, vol. 12, no. 3, pp. 191–201, 2013.

[3] J. J. de Pablo, N. E. Jackson, M. A. Webb, L. Q. Chen, J. E. Moore, D. Morgan, R. Jacobs, T. Pollock, D. G. Schlom, and E. S. Toberer, "New frontiers for the materials genome initiative," *npj Computational Materials*, vol. 5, no. 1, Art. no. 41, 2019.

[4] Agrawal and A. Choudhary, "Perspective: Materials informatics and big data: Realization of the 'fourth paradigm' of science in materials science," *APL Materials*, vol. 4, no. 5, Art. no. 053208, 2016.

[5] *Materials Genome Initiative for Global Competitiveness*. Washington, DC, USA: National Science and Technology Council, Executive Office of the President, 2011.

[6] P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, and A. J. Norquist, "Machine-learning-assisted materials discovery using failed experiments," *Nature*, vol. 533, no. 7601, pp. 73–76, 2016.

[7] L. Ward, A. Dunn, A. Faghaninia, N. E. R. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, K. Chard, M. Asta, K. A. Persson, G. J. Snyder, I. Foster, and A. Jain, "Matminer: An open source toolkit for materials data mining," *Computational Materials Science*, vol. 152, pp. 60–69, 2018.

[8] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, "Machine learning in materials informatics: Recent applications and prospects," *npj Computational Materials*, vol. 3, no. 1, Art. no. 54, 2017.

[9] Y. Liu, T. Zhao, W. Ju, and S. Shi, "Materials discovery and design using machine learning," *Journal of Materiomics*, vol. 3, no. 3, pp. 159–177, 2017.

[10] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent advances and applications of machine learning in solid-state materials science," *npj Computational Materials*, vol. 5, no. 1, Art. no. 83, 2019.

[11] D. M. Dimiduk, E. A. Holm, and S. R. Niezgoda, "Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering," *Integrating Materials and Manufacturing Innovation*, vol. 7, no. 3, pp. 157–172, 2018.

[12] H. K. D. H. Bhadeshia, R. C. Dimitriu, S. Forsik, J. H. Pak, and J. H. Ryu, "Performance of neural networks in materials science," *Materials Science and Technology*, vol. 25, no. 4, pp. 504–510, 2009.

[13] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[14] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD)*, San Francisco, CA, USA, 2016, pp. 785–794.

[15] T. Lookman, P. V. Balachandran, D. Xue, and R. Yuan, "Active learning in materials science with emphasis on adaptive sampling using

uncertainties for targeted design," *npj Computational Materials*, vol. 5, no. 1, Art. no. 21, 2019.

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[17] D. B. Miracle and O. N. Senkov, "A critical review of high entropy alloys and related concepts," *Acta Materialia*, vol. 122, pp. 448–511, 2017.

[18] S. R. Kalidindi and M. De Graef, "Materials data science: Current status and future outlook," *Annual Review of Materials Research*, vol. 45, pp. 171–193, 2015.

[19] J. Wei, X. Chu, X. Y. Sun, K. Xu, H. X. Deng, J. Chen, Z. Wei, and M. Lei, "Machine learning in materials science," *InfoMat*, vol. 1, no. 3, pp. 338–358, 2019.

[20] J. Xiong, T. Y. Zhang, and S. Q. Shi, "Machine learning of mechanical properties of steels," *Science China Technological Sciences*, vol. 63, no. 7, pp. 1247–1255, 2020.

[21] X. Ma, J. Zhao, W. Du, X. Zhang, L. Jiang, and Z. Jiang, "An analysis of dendrite morphology of hypoeutectic Al–Si alloys: The effect of silicon content using machine learning methods," *Computational Materials Science*, vol. 173, Art. no. 109397, 2020.

[22] Meredig, E. Antono, C. Church, M. Hutchinson, J. Ling, S. Paradiso, B. Blaiszik, I. Foster, B. Gibbons, J. Hattrick-Simpers, A. Mehta, and L. Ward, "Can machine learning identify the next high-temperature superconductor? Examining extrapolation performance for materials discovery," *Molecular Systems Design & Engineering*, vol. 3, no. 5, pp. 819–825, 2018.

[23] W. Huang, P. Martin, and H. L. Zhuang, "Machine-learning phase prediction of high-entropy alloys," *Acta Materialia*, vol. 169, pp. 225–236, 2019.

[24] Y. Zhang, C. Wen, C. Wang, S. Antonov, D. Xue, Y. Bai, and Y. Su, "Phase prediction in high entropy alloys with a rational selection of materials descriptors and machine learning models," *Acta Materialia*, vol. 185, pp. 528–539, 2020.

[25] Agrawal and A. Choudhary, "An online tool for predicting fatigue strength of steel alloys based on ensemble data mining," *International Journal of Fatigue*, vol. 113, pp. 389–400, 2018.

[26] R. J. Fruehan, *The Making, Shaping and Treating of Steel: Steelmaking and Refining Volume*, 11th ed. Pittsburgh, PA, USA: AISE Steel Foundation, 1998.

[27] M. Yuan, P. Zhou, H. Wang, and T. Chai, "Prediction of silicon content in hot metal based on data-driven method and mechanism model," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 9, pp. 7887–7896, 2020.

[28] L. Tang, J. Liu, A. Rong, and Z. Yang, "A review on machine learning models for the BOF steelmaking process," *Ironmaking & Steelmaking*, vol. 48, no. 7, pp. 845–861, 2021.

[29] X. Zhang, B. Chen, J. Zhao, and T. Shi, "Real-time detection of breakout for continuous casting using deep learning," *Metals*, vol. 11, no. 2, Art. no. 260, 2021.

[30] G. B. Olson and C. J. Kuehmann, "Materials genomics: From CALPHAD to flight," *Scripta Materialia*, vol. 70, pp. 25–30, 2014.

[31] T. Xie and J. C. Grossman, "Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties," *Physical Review Letters*, vol. 120, no. 14, Art. no. 145301, 2018.

[32] Z. Zhou, Y. Zhou, Q. He, Z. Ding, F. Li, and Y. Yang, "Machine learning guided appraisal and exploration of phase design principles for high entropy alloys," *npj Computational Materials*, vol. 5, no. 1, Art. no. 128, 2019.

[33] J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Science Advances*, vol. 4, no. 6, Art. no. eaar4206, 2018.

[34] Z. Liu and H. K. D. H. Bhadeshia, "Estimation of austenite grain boundaries using thermodynamic models and neural networks," *Metallurgical and Materials Transactions A*, vol. 40, no. 2, pp. 440–449, 2018.

[35] NACE International, *IMPACT—International Measures of Prevention, Application, and Economics of Corrosion Technologies Study*. Houston, TX, USA: NACE International, 2016.

[36] W. Sun, A. Huang, and J. L. Luo, "Predicting corrosion behavior using machine learning: Applications to passive and active corrosion mechanisms," *Electrochimica Acta*, vol. 336, Art. no. 135682, 2020.

[37] X. Wang, T. Shi, G. Tang, Y. Fan, and Z. Hu, "A machine learning approach to predict the atmospheric corrosion of low alloy steels," *npj Materials Degradation*, vol. 5, no. 1, Art. no. 7, 2021.

[38] P. Jain, B. N. Bhatt, and A. Dwivedi, "Neural network-based prediction of hydrogen embrittlement in high-strength steels," *Materials Today: Proceedings*, vol. 48, pp. 840–847, 2022.

[39] M. Popova, O. Isayev, and A. Tropsha, "Deep reinforcement learning for de novo drug design," *Science Advances*, vol. 4, no. 7, Art. no. eaap7885, 2018.

[40] Z. Wu, G. Niu, J. Yang, and M. Chen, "Machine-learning-based predictive model for heat treatment process of automotive medium carbon steels," *Journal of Materials Processing Technology*, vol. 282, Art. no. 116695, 2020.

[41] T. DebRoy, H. L. Wei, J. S. Zuback, T. Mukherjee, J. W. Elmer, J. O. Milewski, A. M. Beese, A. Wilson-Heid, A. De, and W. Zhang, "Additive manufacturing of metallic components – Process, structure and properties," *Progress in Materials Science*, vol. 92, pp. 112–224, 2018.

[42] G. Tapia, S. Khairallah, M. Matthews, W. E. King, and A. Elwany, "Gaussian process-based surrogate modeling framework for process planning in laser powder-bed fusion additive manufacturing of 316L stainless steel," *International Journal of Advanced Manufacturing Technology*, vol. 94, nos. 9–12, pp. 3591–3603, 2018.

[43] Ye, G. S. Hong, Y. Zhang, K. Zhu, and J. Y. H. Fuh, "Defect detection in selective laser melting technology by acoustic signals with deep belief networks," *International Journal of Advanced Manufacturing Technology*, vol. 96, nos. 5–8, pp. 2791–2801, 2018.

[44] A. Holm, R. Cohn, N. Gao, A. R. Kitahara, T. P. Matson, B. Lei, and S. R. Yarasi, "Overview: Computer vision and machine learning for microstructural characterization and analysis," *Metallurgical and Materials Transactions A*, vol. 51, no. 12, pp. 5985–5999, 2020.

[45] B. L. DeCost and E. A. Holm, "A computer vision approach for automated analysis and classification of microstructural image data," *Computational Materials Science*, vol. 110, pp. 126–133, 2015.

[46] S. M. Azimi, D. Britz, M. Engstler, M. Fritz, and F. Mücklich, "Advanced steel microstructural classification by deep learning methods," *Scientific Reports*, vol. 8, no. 1, Art. no. 2128, 2018.

[47] W. Li, K. G. Field, and D. Morgan, "Automated defect analysis in electron microscopic images," *npj Computational Materials*, vol. 4, no. 1, Art. no. 36, 2018.

[48] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, pp. 4765–4774, 2017.

[49] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[50] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in *Proc. 33rd Int. Conf. Machine Learning (ICML)*, vol. 48, pp. 1050–1059, 2016.

[51] N. Walker, A. K. Cheetham, C. D. Sherrill, and A. Aspuru-Guzik, "MatBERT: A materials-aware pre-trained language model," *Nature Machine Intelligence*, vol. 3, no. 8, pp. 665–666, 2021.